DOCUMENT RESUME

ED 346 110 TM 018 326

AUTHOR Edwards, Lynne K.; Bland, Patricia C.
TITLE Some Computer Programs for Selected Problems in Analysis of Variance.
SPONS AGENCY Minnesota Supercomputer Inst.
PUB DATE Apr 92
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Variance; *Computer Software; Computer Software Evaluation; Computer Software Selection; Computer Uses in Education; *Estimation (Mathematics); Generalizability Theory; Mathematical Models; Microcomputers; *Problem Solving; Research Design; *Research Methodology
IDENTIFIERS BMDP (Computer Software); Factorial Designs; Hierarchical Analysis; Statistical Analysis System; Statistical Package for the Social Sciences; Variance (Statistical)

ABSTRACT

Selected examples using the statistical packages Statistical Package for the Social Sciences (SPSS), the Statistical Analysis System (SAS), and BMDP are presented to facilitate their use and encourage appropriate uses in: (1) a hierarchical design; (2) a confounded factorial design; and (3) variance component estimation procedures. To illustrate the hierarchical design, an example is taken from a study of time to assemble an electronic component for five operators at each of six workplaces (Kirk, 1982). The confounded factorial design is illustrated for four subjects each from two independent groups in two treatment combinations (Kirk, 1982). A variance component estimation example is given for a generalizability study involving 10 persons observed on 2 occasions (R. L. Brennan, 1983). For each of these conditions, all three computer packages are used. Nine tables present information from these analyses, and there is a six-item list of references. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
LYNNE K. EDWARDS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED346110

Some Computer Programs for Selected Problems in Analysis of Variance

Paper presented at the 1992 AERA Symposium
in San Francisco

Lynne K. Edwards and **Patricia C. Bland**

University of Minnesota

323 Burton Hall
178 Pillsbury Dr SE
Minneapolis, MN 55455

The funding from the Minnesota Supercomputer Institute is duly acknowledged.

TM018326

BEST COPY AVAILABLE

Introduction

In encouraging researchers to consider, when appropriate, certain designs which are not frequently used, providing an easy access to one of the major statistical packages is a must. Designs such as Latin square and confounded factorial designs, are underemphasized in class partially due to the instructor's unfamiliarity with these designs in the context of educational and psychological research. It is not well understood when these designs are appropriate, nor how they can be analyzed by the available statistical packages.

On the other hand, for some of the more familiar and realistic designs, such as hierarchical or nested designs, specifying the computer commands and to obtain correct results are not as straightforward as we would like. Yet another situation may call for measurement issues such as the generalizability study to be tackled by variance component estimation. When estimated variance components are non-standard, i.e., negative, specifying a default procedure is hardly appropriate.

Selected examples using major statistical packages, SPSS (SPSS, 1990), SAS (SAS, 1989), and BMDP (BMDP, 1988) are presented in this paper in order to encourage appropriate uses of a hierarchical design, a confounded factorial design, and variance component estimation procedures.

1. A Hierarchical Design

Although hierarchical designs involving only a completely randomized structure is a simple design, obtaining correct F values and specifying the data structure have been relatively difficult, if not confusing. A numerical example is used to illustrate the handling of a nested analysis by the three packages, SPSS, SAS, and BMDP. The example is from Kirk (1982, p.485) where an industrial psychologist was interested in decreasing time to assemble an electronic component. Three assembly fixtures, treatment FX, including the one currently in use were evaluated. Five operators at each of six workplaces within the plant, Factor OP(FX), were selected randomly. The assembly fixtures were randomly assigned to the six workplaces and the operators at those workplaces. After a three-week familiarization period, the following data on the number of units assembled per hour were collected.

1.1 Application with SPSS

The data are to be read in four columns: TIME, FX (fixture), WORKPLAC (workplace), and OPERATOR. The data matrix appropriate for either SPSS or SAS readable is presented in Table 1. To handle this problem with SPSS, we have to "fudge" the nesting of WORKPLAC in FIXTURE to make it appear crossed to SPSS. That is, WORKPLAC is recoded to appear as if there were two levels of crossed factor, WP.

```
RECODE      WORKPLAC
            (1,3,5= 1)
            (2,4,6= 2)
            INTO
            WP
```

In the following specification of the dependent and independent variables in the MANOVA procedure, the original six-level WORKPLAC is isolated as a source of variation to facilitate the nested treatment comparisons in the second DESIGN statement. In this context, note that the CONTRAST works in conjunction with the DESIGN statement immediately following it.

```
MANOVA      TIME by WORKPLAC(1,6), FX(1,3), WP(1,2)/
            DESIGN= FX VS 1,
                    WP WITHIN FX=1 VS WITHIN/
            CONTRAST (WORKPLAC)= SPECIAL
                           (  1   1   1   1   1   1
                              1   1   0   0  -1  -1
                              0   0   1   1  -1  -1
                              1  -1   0   0   0   0
                              0   0   1  -1   0   0
                              0   0   0   0   1  -1)/
            PARTITION(WORKPLAC)= (2, 1, 1, 1)/
            DESIGN=WORKPLAC (2) VS WITHIN,
                   WORKPLAC (3) VS WITHIN,
                   WORKPLAC (4) VS WITHIN/
```

The first DESIGN statement specifies that FX is to be tested against Error1 and Workplace nested in FX is to be tested against within cells. The contrast specifications define Grand Mean in the first row, which is automatically taken out of the total, FX with two d.f. in the 2nd and 3rd rows, a comparison of workplaces at each level of FX is in the 4th, 5th, and 6th row. The partition statement specifies that the SSworkplac is to be decomposed into four separate SSs: the 2nd and 3rd rows combined representing FX with 2 d.f. and a 1 d.f. each for the nested analysis at each level of FX. Recall that the first contrast is automatically excluded as it represents the grand mean.
The overall test as well as a followup analysis at each workplace are presented in Table 2. The design statement specifies that the error term for the last three nested analyses is the within cells.

1.2 Applications with SAS

As in SPSS, SAS GLM procedure reads in four columns of data, TIME, FX, WP, and OP for Operator. First PROC GLM performs the overall ANOVA using the least squares regression approach, SS3 (i.e., SS type III), followed by specific hypothesis tests [i.e., TEST H=FX E=WP(FX), and TEST H=WP(FX) E=OP(WP FX)]. The error for the second hypothesis test is OP(WP(FX)), the within cells variation. However, there is no double-parenthesis format in the

SAS commands, thus, it is presented as OP(WP FX). For hypothesis tests, the error term must be explicitly specified, thus OP(WP FX) in the model specification. If hypothesis tests using different error terms were not desired, OP(WP FX) would simply be reported as ERROR by SAS, and all F-tests and contrasts would "default" to that source of variation.

```
PROC GLM;
     CLASS FX WP OP;
     MODEL TIME=FX WP(FX) OP(WP FX)/SS3;
     TEST H=FX E=WP(FX);
     TEST H=WP(FX) E=OP(WP FX);
```

The next segment specifies followup nested analyses by specifying contrast coefficients.

```
PROC GLM;
     CLASS FX WP;
     MODEL TIME=FX WP(FX)/SS3;
     CONTRAST 'FIXTURE 1, WP 1 VS 2'
          WP(FX) 1 -1  0  0  0  0;
     CONTRAST 'FIXTURE 2, WP 3 VS 4'
          WP(FX) 0  0  1 -1  0  0;
     CONTRAST 'FIXTURE 3, WP 5 VS 6'
          WP(FX) 0  0  0  0  1 -1;
```

The result for the second segment of PROC GLM is in Table 3. The error term for FX erroneously uses the default of Within Cells. The correct result for FX is obtained in the first segment of PROC GLM, however, the printout is slightly awkward.

1.3 Applications with BMDP

BMDP8V uses the statistical model and its indices to organize data for input. The way in which the data are organized is critical with respect to the order of the indices. To see how these data correspond to an "observation-wise" structure, see the corresponding SPSS example. Specific contrasts, as shown in the SPSS and SAS examples, require a reorganization of the data into an "observation-wise" structure.

BMDP8V requires that the data for this example to be read in six columns, each column representing five observations taken from a particular fixture at a particular workplace, e.g., Fix1Wp1 is a column of observations taken at Fixture 1 at Workplace 1 as indicated below.

17	13	21	25	26	32
7	23	24	29	34	32
9	14	14	26	32	37
12	18	19	24	30	34
15	22	17	21	28	35

The desirable feature in the BMDP8V is that a user can specify whether the factors are fixed or random. In addition, the statistical model including the nested structure can be specified without needing to instruct the program which error term to use.

```
/DESIGN   DEPENDENT=Fix1Wp1 to Fix3Wp6.
          LEVELS=5, 3, 2.
          NAMES= OP, FIXTURE, WP.
          FIXED= FIXTURE.
          RANDOM= OP, WP.
          MODEL= 'F, W(F), O(FW)'.
```

The tabled ANOVA summary table for the overall tests are presented in Table 4. It has the clearest form of presentation with each F ratio presented with the corresponding appropriate mean squares. Unfortunately, the followup nested analyses have to be handled separately in BMD7D, and the results are not reported here.

2. A Confounded Factorial Design

Various forms of confounded factorial designs are presented in Kirk (1982), however, a confounded factorial design remains one of the least utilized designs in education and psychology. Devised as a solution to a large and impractical block-size problem, only a portion of the treatment combinations is assigned to each block in confounded factorial designs. The block size is often reduced by way of confounding group with interaction. A 2^2 randomized block confounded design with the block size of two (RBCF-2^2) can be compared to a 2x2 randomized block factorial design with the block size of four (RBF-22). If it is possible to assign "matched" four subjects to a block size of four, the advantage of a confounded factorial design over a randomized block factorial design is not obvious. However, when the block size is constrained to two but the number of treatment combinations exceeds two, such as the case in a twin study where three or more treatment combinations are examined, a confounded factorial design becomes an viable option.

An example RBCF-2^2 is taken from Kirk (1982, p.580), four subjects each from two independent groups underwent two different treatment combinations. The first level of each factor is indicated by 0, as in g0 for the first group, and the kth level by k-1. The first group (g0) were measured under a0b0 and a1b1 combinations and the second group (g1) were measured under a0b1 and a1b0 combinations.

	ab 00	ab 11
s0	3	16
s1	5	14
s2	6	17
s3	5	15
	01	**10**
s4	14	7
s5	14	6
s6	16	7
s7	16	11

2.1 Applications with SPSS

The data contains five columns: Group factor level, A factor level, B factor level, Block level, then Y observation (see Table 5). Each factor's first level is set to 0, the second to 1, and the kth to k-1. In this example, A and B factor combinations are first translated into COLUMN variable. That is, whether the responses in question appear in Column 1 (COLUMN=1) or in Column 2 (COLUMN=2) is noted. This manipulation is not necessary, but it makes the point that because a fraction of the treatment combinations are used, a comparison of two columns is identical to a comparison of either A or B factor.

```
IF (A EQ 0 AND B EQ 0) COLUMN= 1
IF (A EQ 0 AND B EQ 1) COLUMN= 1
IF (A EQ 1 AND B EQ 1) COLUMN= 2
IF (A EQ 1 AND B EQ 0) COLUMN= 2
```

Because SPSS requires that all effects to appear crossed, we "fudge" the nesting of BLOCKS in GROUPS to make it appear crossed to SPSS.

```
RECODE     BLOCK
           (0=0) (1=1) (2=2) (3=3) (4=0) (5=1) (6=2) (7=3)
           INTO
           BL
```

First, the homogeneity of MSABxBL(G0) and MSABxBL(G1) is examined. If homogeneous, they can be pooled to form the MSABxBL. The following SPSS commands produces these two error estimates then the summary table using the pooled error term (Table 6).

```
TEMPORARY
SELECO IF (G EQ 0)
MANOVA     Y BY COLUMN(1,2), BL(0,3)/
           DESIGN=COLUMN, BL/
TEMPORARY
```

```
SELECT IF (G EQ 1)
MANOVA     Y BY COLUMN(1,2), BL(0,3)/
           DESIGN=COLUMN, BL/
MANOVA     Y BY G, A (0,1), B(0,1), BL(0,3)/
           DESIGN=G VS 1,
                BL WITHIN G=1,
                A VS RESIDUAL,
                B VS RESIDUAL/
```

Since the residual variances from G0 and G1 were homogeneous (identical in this case), the pooled residual is used in the last MANOVA procedure, where the pooled residual is used as the error for testing A and B effects and the error 1 [BL within G] is used to test Group effects.

2.2 Applications with SAS

SAS has a simple command structure in obtaining Group 0 and Group 1 analyses and the overall analysis. The data structure is the same as in SPSS. The first two GLM procedures produce the residual variance for G0 and G1 as the error variance for the model and the third GLM produces the summary table using the pooled residual. In the overall summary table, one incorrect F ratio is reported in testing G against BL(G) (see Table 7). A and B are tested against the error from the model, MSerror = 1.458. In order to obtain the correct tet of G, a specific test is requested using BL(G), MSBL(G) = 3.292, as the denominator term in testing G. The SAS result for the first two segments showing the error variances from the models with G0 and G1 is not presented in Table 5.

```
PROC GLM;
          WHERE G=0;
          CLASS     G  A  B  BL;
          MODEL  Y=A BL;
PROC GLM;
          WHERE G=1;
          CLASS     G  A  B  BL;
          MODEL  Y=A BL;
PROC GLM;
          CLASS     G  A  B  BL;
          MODEL  Y=A  B  G  BL(G);
          TEST   H=G     E=BL(G);
```

2.3 Applications with BMDP

Confounded factorial designs are not well handled in BMDP4V. The data structure is the same as for the ones used in SAS and SPSS.

```
/BETWEEN  FACTORS ARE G, A, B, BLOCK.
/GROUP    CODES(G) = 0, 1.
          CODES(A) = 0, 1.
          CODES(B) = 0, 1.
```

```
          CODES(BLOCK)=0, 1, 2, 3, 4, 5, 6, 7.
/WEIGHTS  BETWEEN=EQUAL.
/END
     ...Data are read in
/END
ANALYSIS  PROCEDURE=STRUCTURE.
          BFORMULA= 'A+B+G+BLOCK.G'./
```

The first analysis procedure produces the table with all correct tests except for the test of Groups. In order to produce the appropriate test for G, we resort to specifying a contrast matrix for BLOCK, which is cumbersome.

```
DESIGN    FACTORS=A, G, BLOCK.
          TYPE =BETWEEN, CONTRAST.
          CODE=READ.
          VALUES ARE
           1, 0, 0, 0, 0, 0,
           0, 1, 0, 0, 0, 0,
           0, 0, 1, 0, 0, 0,
          -1,-1,-1, 0, 0, 0,
           .  .  .  .  .  .
           .  .  .  .  .  .
           .  .  .  .  .  .
           0, 0, 0, 1, 1, 1./
ANALYSIS  PROCEDURE=STRUCTURE.
          BFORMULA='A+B+G'./
```

This arrangement isolates the Residual as a source of variation, thus, everything else is tested against a new error which is left after extracting A, B, G, and that is labeled as AGB. The output is not presented but the incorrect test of G in the first analysis is analogous to the SAS output without the TEST specification. Of all three packages examined, BMDP is the most inflexible and awkward program for handling confounded factorial designs.

3. A Variance Component Estimation Example in Measurement

An example of a generalizability study from Brennan (1983, p. 47) is used to illustrate variance component estimation using a computer package. In the example 10 persons, P, were observed under two occasions, O, with different sets of four Items nested in occasion, I(O). Because P is crossed with O, the effects of interests are: P, O, I(O), PO, PI(O). The data are presented in Table 8. This example is similar to that of a repeated measures design, where each subject is measured under 2x4 = 8 treament combinations.

The variance component estimation is not properly handled in SPSS. BMDP3V has a single-precision result with three decimal places for either ML or REML. Therefore, only a viable option is SAS VARCOMP. Thus, only the example from SAS VARCOMP is used for

this illustration.

First, because SAS handles data as crossed, recoding is necessary to let the system know that Item 1 at Occasion 2 is equal to Item 5, and so forth.

```
IF   O=2  AND I=1   THEN I=5;
IF   O=2  AND I=2   THEN I=6;
IF   O=2  AND I=3   THEN I=7;
IF   O=2  AND I=4   THEN I=8;
```

After the data are read in PROC VARCOMP is specified. The choices of methods are: TYPE1, which uses Type I SS or labeled as AOV(I) estimation; MIVQUE0, Minimum variance quadratic estimator with no iteration; ML, Maximum Likelihood estimation with iterations; and REML, Restricted Maximum Likelihood estimation with iterations. The first two estimations may report negative variances, but the ML and REML in VARCOMP automatically replaces negative estimates with 0s, thus introducing some bias. [For more information on different estimation procedures, see Hocking (1992).]

```
PROC VARCOMP   METHOD= REML;
     CLASS     PERSON    O    I;
               MODEL     Y=PERSON
                           O
                           I(O)
                           PERSON*O;
```

Note that the correct error term is PERSON*I(O), however, if this term is specified in the model, SAS will produce its own error with expectation of 0 and its d.f. of 0. In the example provided, this causes little difference in MVQUE0, ML, and REML solutions, however, SAS refuses to decompose variances in TYPE1 method. [For the two iterative methods, ML and REML, the over-specified model produced solutions with fewer iterations than the model leaving the error.]

TYPE1 procedure is what Hocking named AOV(I), Anova solution using type I SS, and is solved on the basis of the expected mean squares listed in Table 9. In Brennan's (1983) terminology, TYPE1 is Algorithm 2 presented in his Table 3.4.1 (p. 47). A comparison of all four procedures, TYPE1, MVQUE0, ML, and REML is in Table 10.

In this balanced example, that is, without any missing value, the negative component estimates may be due to the sampling variability, in particular, from a small sample as Brennan suggests. However, outliers or unusual observations also can result in negative estimates. For balanced and positive estimates, the simplest procedure TYPE1 is recommended. But once negative estimates are obtained, Hocking recommends REML for its generally fewer iterations and smaller bias in comparison to the ML procedure. Yet another possibility is a misspecification of the model. Hocking recommends a diagnosis of the data when faced

with negative estimates. For example, a scatter plot can be prepared to examine possible extreme values in the data. In this example, Subject 9 at Occasion 1, Item 1 with 0 score may be slightly off from the general cluster of data points. Because the data used are artificial in this example, it is difficult to ascertain the cause of negative estimates but Hocking's recommendation for understanding the sources of negative estimates is warranted. Replacing negative estimates with zeros is a simply a fix without telling us the nature of data, but we should use the information as a guide to a full diagnostic of data.

In this example, TRYPE1 (and MVQUE0), ML, and REML estimates are different. In non-iterative methods, TYPE1 and MVQUE0, two negative estimates are reported. Since some variance components are negative, TYPE1 or MVQUE0 solutions are not recommended. Because of a larger bias in ML than in REML, it is recommended to choose REML estimates, we then would allocate noticeably smaller variance to Items nested in Occasion than we would have with TYPE1 estimates.

Table 1. A Hierarchical ANOVA Example Data for SPSS

TIME	FX	WORKPLAC	OPERATOR
17	1	1	1
7	1	1	2
9	1	1	3
12	1	1	4
15	1	1	5
13	1	2	6
23	1	2	7
14	1	2	8
18	1	2	9
22	1	2	10
21	2	3	11
24	2	3	12
14	2	3	13
19	2	3	14
17	2	3	15
25	2	4	16
29	2	4	17
26	2	4	18
24	2	4	19
21	2	4	20
26	3	5	21
34	3	5	22
32	3	5	23
30	3	5	24
28	3	5	25
32	3	6	26
32	3	6	27
37	3	6	28
34	3	6	29
35	3	6	30

Table 2. Hierarchical ANOVA and Nested Treatment Comparisons with SPSS

ANOVA summary table

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN CELLS	300.00	24	12.50		
WP WITHIN FX (Error 1)	220.00	3	73.33	5.87	.004
Error 1	220.00	3	73.33		
FX	1460.00	2	730.00	9.95	.047
Nested Treatment Comparisons					
WITHIN CELLS	300.00	24	12.50		
WORKPLAC(2)	90.00	1	90.00	7.20	.013
WORKPLAC(3)	90.00	1	90.00	7.20	.013
WORKPLAC(4)	40.00	1	40.00	3.20	.086

Table 3. SAS result for the hierarchical data

Source	DF	Type III SS	Mean Square	Fvalue	Pr<F
First GLM segment					
FX	2	1460.00000000	730.00000000	.	.
WP(FX)	3	220.00000000	73.33333333	.	.
OP(FX*WP)	24	300.00000000	12.50000000	.	.
Tests of Hypotheses using the Type III MS for WP(FX) as an error					
FX	2	1460.0000	730.00000000	9.95	.0474
Tests of Hypotheses using the Type III MS for OP(FX*WP) as an error					
WP(FX)	3	220.0000	73.33333333	5.87	.0037
Second GLM segment					
FX	2	1460.00000000	730.00000000	58.40	0.0001
WP(FX)	3	220.00000000	73.33333333	5.87	0.0037
Contrast					
Fix1, wp1 vs 2	1	90.00000000	90.0000000	7.20	0.0130
Fix2, wp3 vs 4	1	90.00000000	90.0000000	7.20	0.0130
Fix3, wp5 vs 6	1	40.00000000	40.0000000	3.20	0.0863

Table 4. Hierarchical ANOVA with BMDP8V

Source	Error term	SS	d.f.	MS	F	PROB
MEAN	W(F)	15870.000	1	15870.00	216.41	.0007
FIXTURE	W(F)	1460.000	2	730.00	9.95	.0474
W(F)	O(FW)	220.000	3	73.33	5.87	.0037
O(FW)		300.000	24	12.50		

Table 5. A Data Matrix for a Confounded Factorial Example

G	A	B	BLOCK	Y
0	0	0	0	3
0	1	1	0	16
0	0	0	1	5
0	1	1	1	14
0	0	0	2	6
0	1	1	2	17
0	0	0	3	5
0	1	1	3	15
1	0	1	4	14
1	1	0	4	7
1	0	1	5	14
1	1	0	5	6
1	0	1	6	16
1	1	0	6	7
1	0	1	7	16
1	1	0	7	11

Table 6. A Confounded Factorial 2^2 Design SPSS Summary Table

Source of Variation	SS	DF	MS	F	Sig of F
G0					
RESIDUAL	4.38	3	1.46		
COLUMN	231.12	1	231.12	158.49	.001
BL	5.37	3	1.79	1.23	.436
G1					
RESIDUAL	4.38	3	1.46		
COLUMN	105.13	1	105.13	72.09	.003
BL	14.37	3	4.79	3.29	.177
RESIDUAL	8.75	6	1.46		
A	12.25	1	12.25	8.40	.027
B	324.00	1	324.00	222.17	.000
Error 1	19.75	6	3.29		
G	6.25	1	6.25	1.90	.217

Table 7. A Confounded Factorial SAS Summary Table

Source	DF	Type III SS	Mean Square	Fvalue	Pr>F
A	1	12.25000000	12.25000000	8.40	.0274
B	1	324.00000000	324.00000000	222.17	.0001
G	1	6.25000000	6.25000000	4.29	.0839
BL(G)	6	19.75000000	3.29166667	2.26	.1725

Test of Hypotheses using the Type III MS for BL(G) as an error

Source	DF	Type III SS	Mean Square	Fvalue	Pr>F
G	1	6.25000000	6.25000000	1.90	.2174

Note: The test for G appearing with the F value of 4.29 is incorrect. The correct test of G is performed on the bottom where G is tested against BL(G).

Table 8. SAS Data for a Variance Component Estimation Example

P	O	I	Y	continued	P	O	I	Y
1	1	1	2		1	2	1	2
2	1	1	4		2	2	1	6
3	1	1	5		3	2	1	5
4	1	1	5		4	2	1	5
5	1	1	4		5	2	1	4
6	1	1	4		6	2	1	6
7	1	1	2		7	2	1	2
8	1	1	3		8	2	1	6
9	1	1	0		9	2	1	5
10	1	1	6		10	2	1	6
1	1	2	6		1	2	2	5
2	1	2	6		2	2	2	7
3	1	2	5		3	2	2	4
4	1	2	9		4	2	2	7
5	1	2	3		5	2	2	5
6	1	2	4		6	2	2	4
7	1	2	6		7	2	2	7
8	1	2	4		8	2	2	6
9	1	2	5		9	2	2	5
10	1	2	8		10	2	2	8
1	1	3	7		1	2	3	5
2	1	3	6		2	2	3	5
3	1	3	4		3	2	3	5
4	1	3	8		4	2	3	7
5	1	3	5		5	2	3	6
6	1	3	4		6	2	3	7
7	1	3	6		7	2	3	7
8	1	3	4		8	2	3	6
9	1	3	4		9	2	3	5
10	1	3	7		10	2	3	8
1	1	4	5		1	2	4	5
2	1	4	7		2	2	4	7
3	1	4	6		3	2	4	5
4	1	4	6		4	2	4	6
5	1	4	6		5	2	4	4
6	1	4	7		6	2	4	8
7	1	4	5		7	2	4	5
8	1	4	5		8	2	4	4
9	1	4	5		9	2	4	3
10	1	4	6		10	2	4	6

continues to right

Table 9. Expected Mean Square for TYPE1 Estimation

Source	Expected Mean Square
PERSON	Var(Error) + 4 Var(PERSON*O) + 8 Var(PERSON)
O	Var(Error) + 4 Var(PERSON*O) + 10 Var(I(O)) + 40 Var(O)
I(O)	Var(Error) + 10 Var(I(O))
PERSON*O	Var(Error) + 4 Var(PERSON*O)
Error	Var(Error)

Table 10. A Variance Components Estimation Example with Negative Estimates

Variance Components	Type1[a]	ML[b]	REML[c]
Var(PERSON)	.69652778	.62624163	.67817628
Var(O)	-.10300926	0	0
Var(I(O))	.59814815	.48804816	.53928705
Var(PERSON*O)	-.04282407	0	0
Var(Error)	1.51018519	1.48880543	1.48571359

a Note the data are balanced, therefore, the results from Type1 and MVQUE0 are identical.
b The number of iterations is 7.
c The number of iterations is 6.

References

Brennan, R. L. (1983). Elements of generalizability theory. Iowa City: ACT Publications.

Dixon, W. J., Brown, M. B., Engelman, L., Hill, M. A., & Jennrich, R. I. (1988a). BMDP statistical software manual (Vol. 1). Berkeley, CA: U of Calif. Press.

Dixon, W. J., Brown, M. B., Engelman, L., Hill, M. A., & Jennrich, R. I. (1988b). BMDP statistical software manual (Vol. 2). Berkeley, CA: U of Calif. Press.

Hocking, R. R. (in press). Variance component estimation in mixed linear models. In L.K. Edwards (ed.), Applied analysis of variance in behavioral sciences. New York: Marcel Dekker.

SPSS Inc. (1990). SPSS reference guide. Chicago, IL: Author.

SAS Institute (1989). SAS user's guide: Statistics (6th ed.). Cary, NC: Author.